AWARD NUMBER: W81XWH-15-2-0016

TITLE: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

PRINCIPAL INVESTIGATOR: Dr. Jesus Caban

CONTRACTING ORGANIZATION: The Geneva Foundation
917 Pacific Ave, Suite 600
Tacoma, WA 98402

REPORT DATE: May 2019

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release; Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.
We propose to design, develop, and validate an interactive visualization framework that physicians assessing TBI/PTSD patients with comorbid symptoms can use to explore and analyze clinical data and that researchers can use to hypothesize new research questions. The primary aims of this project are to (1) extend our interactive visual analytic framework which combines multiple clinical measurements to allow it to be used to explore large collections of clinical data and (2) validate the effectiveness of such visualization systems among clinicians that treat service members.
TABLE OF CONTENTS:

Example Table of Contents:

| 1. Introduction       | 4 |
| 2. Keywords           | 4 |
| 3. Accomplishments    | 4 |
| 4. Impact             | 12 |
| 5. Changes/Problems   | 12 |
| 6. Products           | 13 |
| 7. Participants & Other Collaborating Organizations | 13 |
| 8. Special Reporting Requirements | 14 |
| 9. Appendices         | 17 |
INTRODUCTION: An Interactive Visualization Framework to Support Exploration and Analysis of TBI/PTSD Clinical Data

1. **KEYWORDS:** Data Visualization, Health Information Technologies, TBI/PTSD, Open Source Tools

2. **ACCOMPLISHMENTS:** The overarching goals of this project are to (1) address the gap between the acquisition of clinical measurements and the diagnosis step by providing an intuitive, flexible, and customizable interactive data visualization framework and (2) validate the system among clinicians treating service members diagnosed with TBI / PTSD.

During the first year of the award, a significant amount of work was accomplished. The work accomplished during the first two years was extended to continue to reach each of the tasks and deliverables of the award. The work accomplished during the last year can be summarized as:

1. [Aims 1.7, 1.8, 2.3, 2.4, 2.5] Update system based on usability study and suggestions received from users.

Summary: Despite the growth of visual analytics and the widespread use of visualization tools, the mining and modeling of user interactions and behaviors have not kept pace despite the ability to provide critical information about the method in which users interact with the interface and visualization. Existing approaches contain substantial limitations as they rely on expertly crafted rules or for users to examine their own interactions. As part of Aim 1.8 we introduced the application of reinforcement learning onto user interactions within a visual analytic system, opening a new avenue of research for building models that can automatically learn and uncover hidden insights. We outlined three research directions that can be addressed using reinforcement learning and used to test the techniques developed under this grant: (i) providing guidance to users towards the completion of a task, (ii) personalizing guidance to match each individual user’s behavior and preferences such that the guidance is “polite” and unobtrusive, and (iii) optimizing the visual analytic interface to maximize user performance. Then we present our approach focused on solving these research directions. We show how our approach builds a model that embeds knowledge about a visual analytic system and how it can automatically learn to tailor itself to a user’s preferences through the application of providing guidance. Finally, we manually validate the insights embedded within our model to show how it could be further applied to optimizing an interface. Our reinforcement learning-based models can be utilized with user data or in an automatic standalone method, and ultimately allow for analysis that is not possible using standard modelling techniques.
2. [Aims 1.3, 1.4] Design a modularized visual analytics framework capable of loading, filtering, and illustrating any number of clinical variables.

Summary: Automatically summarizing a collection of temporal sequences is a difficult task given the irregular and variable patterns often found in longitudinal events. Across a wide array of domains, researchers and analysts seek to determine ways to gain an overview of event datasets through the identification of the common paths as well as the trajectories that exist between individual events. While these tasks have been thoroughly researched, many approaches have not taken into account the temporal context of events and thus vital time information is commonly lost. An application has designed to visually explore temporal sequences by incorporating the context of time into event mining algorithms and then overcoming the challenges and complexities of visualizing large scale longitudinal data. The key contributions of this work are an adaptive window-based frequency sequence mining algorithm designed to identify common subsequences and build overviews of longitudinal trajectories and a novel event summary diagram to display common paths while providing multiple levels of details. The algorithm and visualization techniques have been evaluated through three case studies applied to large-scale longitudinal datasets from different domains.
3. Performed an in-depth literature review of existing work related to temporal trajectory analysis that can be used to illustrate clinical data for TBI/PTSD patients. The visualization and analysis of temporal event data is a widely studied topic. In our literature review, we noticed two series of general studies, one that focused on lifelines and another one that was more focused on flow visualization. Table #1 summarizes our findings:
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1:** Summary of papers reviewed related to temporal trajectory analysis.

All scientific papers reviewed were put into specific categories. Papers #1 to #5 were designed for visualization of aggregated data. An important step in these studies was the mining or aggregating the temporal data to find patterns and then employing visualization techniques such as the Sankey diagram to illustrate the longitudinal aspects of the data. Papers #6-#8 were designed for visualization of individual data points. Lifelines was designed for personal patients clinic history visualization, although the other two methods Lifelines2 and PatternFinder are designed for multiple records, they visualize the multiple records separately. Finally, paper #9 (ActiviTree) was a method developed to allow users
to explore the event sequences rather than aggregate patterns form the multiple temporal sequences.

4. [Aims 1.3, 1.7, 1.8] Designed the *Necklace* interface. In order to help users explore the trajectory of medical claim data and compare trajectories and costs, we designed *Necklace*. The Necklace interface consists of three main design ideas: (1) *guided interaction*; (2) *position based global pattern finding*; (3) *multivariate comparison*. See Figures #1 and #2 for more information.

There are two kinds of rings that are part of the Necklace interface: root ring and branch ring. The initial display only shows a root ring, which contains the information of all the transformation and relationships in the dataset. As users interact with the interface, Necklace will show more and more branch rings. Each branch ring has a parent ring and a parent trajectory, and contains a parent event E and the either the incoming events of E or outgoing events of E. Thus a trajectory is composed of the parent events of a series linked rings.

Each ring contains four components: nodes, groups, links and chords. Each node on the ring represents a diagnosis event, and those events on branch rings are not individual events, they present events in a trajectory. For instance, assuming there is a branch ring R which shows its parent event E₃, and those outgoing events of E₃, and the parent trajectory of R is E₁ → E₂ → E₃, so an event E₄ on R represents the event E₄ in the trajectory E₁ → E₂ → E₃ → E₄. And using the ICD 9 Code, we clustered those nodes on a ring into groups. And each link inside the ring represent a transformation relationship between two events, and links from one group to another group are gathered into a chord.

The initial heights of nodes are similar and link nodes encode the incident rate of the corresponding event, which means that the more links, the higher the node will be. In addition, in each node there is a histogram for users to show more information. A ring shows abstract information when it occupies a small area, the detailed nodes information and links information will be hidden if the area is not big enough. In order to obtain detailed nodes information, users can click on outer contour of a group to expand this group to a higher level for more space, and they can also click on the inner contour of a group to shrink to a lower level. One of the advantage of ring is that it has a powerful ability to show large scaled event since a ring can show thousands of event at a time.
Figure 4: Screenshot of the Necklace system. Necklace visualizes temporal events trajectories, this screen shot shows CMS data, which contains 222 patients and 668 claim records. Users can explore the diagnoses trajectory by directly interacting with the visual objects shown in C, and compare trajectories of different patients cohorts with the overview graphs shown in B. And Necklace also offers a user panel, which is shown in A, for users to edit patient cohorts, search for diagnosis node and zoom the display.

Figure 5: Screenshot of the Necklace system. Necklace visualizes temporal events trajectories, this screen shot shows CMS data, which contains 222 patients and 668 claim records.
5. [Aims 1.3, 1.7, 1.8] Developed, extended, and continued to enhance the Patient Timeline tool based on comments and feedback received from different clinical staff members. The Patient Timeline tool was developed to assist physicians explore the longitudinal medical data of a patient. In a visually appealing way, the Patient Timeline starts by displaying a tree that provides a quick glimpse of the data available for a patient. Icons and their size show if the patient has a certain type of data in their history and how much or how important that data is. The year nodes can be expanded to show another level of the tree, showing the patient’s monthly data for that year. To the right of the patient’s tree is a summary panel. This summary panel shows a brief text summary of the patient, their lab history, their diagnosis history and their medication history. The Labs, Diagnoses, and Medications tabs display their information in analytical models that make the data easier to digest. The Labs, Diagnoses, and Medications tabs update when a new tree node is clicked. Furthermore, three patients’ tree can be seen by selecting the Three Trees display from the dropdown button at the top of the page.

![Figure 6: The Patient Timeline starts by displaying a tree that provides a quick glimpse of the data available for a patient.](image)

In the more robust portion of the Patient Timeline, the nodes are presented to us on a timeline as shown in Figure #3. Here, individual days of the patient’s data can be explored in panels. Within the panels there are tabs for Labs, Medications, Vitals, Notes, Diagnoses, Procedures, Radiology Note, and Chief Complaint for a single day, assuming the patient has that type of data available for that day. To the left of the timeline, we have a filter that allows the user to filter nodes based on Provider Type, Provider Name, DMIS/MTF, MEPER4, Encounter Type, Data Type, and/or Date. To the right we have a smaller version of the, previously presented, summary panel.
6. [Aims 1.3, 1.7, 1.8, 1.9] In order to also satisfy the requirements from research clinicians interested in exploring clinical data but also understanding what data is available, we designed the *Database Search* tool. The *Database Search* tool was designed to facilitate the searching and exploration of databases. When the user clicks the Run Query button, the Results panel fills with statistics of their selected tables and variables. In the Results panel, they can see the Count, Missing, Mean, Standard Deviation, Zeroes, Minimum, Maximum, and Histogram for their chosen variables. Once the user has selected all the data they would like to put a request in for, they can click the Export button, followed by the ‘Copy to clipboard’ button to copy their selections to their computer’s clipboard. This saves the user from having to individually copy and paste, or type up, the tables and variables they would like to put a request in for.
Figure 8: Screenshots of the Database Search tool designed to better explore clinical data that is available to research purposes.

7. [Aims 1.3, 1.7, 1.8, 1.9] A new module was prototyped to visually analyze clinical notes. The tool automatically highlights clinical concepts within unstructured provider notes. The new module can be used to compare clinical notes and highlight changes over time. In addition, the new module can be used to identify commonalities between different EHR notes.
8. [Aims 1.3, 1.7, 1.8, 1.9] In order to better understand and validate the effectiveness of different visualization tools, we have developed a method to capture how users interact with the different systems. Given a clinical dataset and a visualization tool or dashboard, we designed an application to compare the correlations, patterns, and flows different groups follow to reach an answer for a specific clinical question.

Figure 9: Screenshots of the module under development to analyze / compare clinical notes.
What were the major goals of the project?

The first major goal of this project was to design a visual analytic framework that combines multiple clinical measurements and allows the exploration of large collections of clinical data. In addition, the second major goal (years 2-4) was to validate the effectiveness and usability of different visualization techniques for exploring large collections of clinical variables with complex associations.

The first major goal of the second year was to iterate on our work for summarizing large, longitudinal datasets by expanding on our previous work and developing a novel visualization method for analyzing a large dataset of events. In addition, we validated the effectiveness of our visualization techniques throughout the year. The second major goal of the second year was to build a system that summarizes the various elements of a patient’s EHR data. Through these two goals, we have addressed the many difficulties associated with clinical variables and data from both a big data and a singular patient-provider perspective.

The first major goal of the third year was to continue to enhance the existing tools from previous years by adding new visualization techniques, new validation approaches, and new data. That was accomplished by continuing to perform a literature review, develop the Necklace interface, continue to enhance the patient time-line framework, develop an effective interface for research clinicians to search at data, and by creating a systematic approach to analyze the patient...
interaction data. During the next year the priority will be to continue to enhance those tools and make them available to other researchers, investigators, and organizations.

The first major goal of the four year was to validate the different applications using standardized techniques. During the final year the priority will be to continue to enhance make applications available to other researchers, investigators, and organizations as well as close out the project.

What was accomplished under these goals?

During the first year, four different systems were prototyped and developed to perform visualization of tabular, hierarchical, and longitudinal data. First, VisXplore was enhanced to become a clinical data visualization system to perform group or single-subject analysis of multivariate tabular, hierarchical, or temporal clinical data. Second, CoFlow was developed as an interactive multi-view and exploratory visualization tool designed to analyze longitudinal EHR data. Third, a graph-based visualization technique was developed to visually explore the frequency of patients going from one specific clinical diagnosis to other diagnosis. Finally, a visual summarization approach was created and tested with thousands of mTBI patients. Each of the tool has a corresponding draft paper describing the design and techniques. See attachments.

During the second year, two different systems were extensively prototyped and developed to effectively summarize the various data elements that are present in Electronic Health Records (EHRs). First, a novel visualization method, event summary diagrams, and a corresponding system were built to enable for a large dataset of events to easily be understood through a top-down interactive exploration. This visualization was evaluated with a dataset of thousands of mTBI patients and shown to reduce the visual complexity and analytical capacity required compared to existing techniques. Second, a timeline-based framework for aggregating and summarizing EHRs was extensively researched, designed, and developed to overcome the challenges that exist in EHR systems where data integration is lacking and the disparate nature of data creates difficulties for clinicians. Through this framework, a clinician is able to view the entire history of a patient at multiple time scales and develop an understanding of the patient state over time. Each of these tools have a corresponding draft paper describing the design and techniques. See attachments.

During the third we continued to perform a literature review, developed the Necklace interface, continued to enhance the patient time-line framework, developed an effective interface for research clinicians to search at data, and by creating a systematic approach to analyze the patient interaction data. During the next year the priority will be to continue to enhance those tools and make them available to other researchers, investigators, and organizations.

During the four year we introduced the application of reinforcement learning onto user interactions within a visual analytic system, opening a new avenue of research for building models that can automatically learn and uncover hidden insights. We outlined three research directions that can be addressed using reinforcement learning and used to test the techniques developed under this grant: (i) providing guidance to users towards the completion of a task, (ii)
personalizing guidance to match each individual user’s behavior and preferences such that the
guidance is “polite” and unobtrusive, and (iii) optimizing the visual analytic interface to
maximize user performance. Then we present our approach focused on solving these research
directions. We show how our approach builds a model that embeds knowledge about a visual
analytic system and how it can automatically learn to tailor itself to a user’s preferences through
the application of providing guidance. Finally, we manually validate the insights embedded
within our model to show how it could be further applied to optimizing an interface. Our
reinforcement learning-based models can be utilized with user data or in an automatic standalone
method, and ultimately allow for analysis that is not possible using standard modelling
techniques.

What opportunities for training and professional development has the project provided?

“Nothing to Report.”

How were the results disseminated to communities of interest?

Some of the prototype visualization tools were demonstrated during the 2018 Workshop on
Visual Analytics in Healthcare and at the 2018 IEEE Visualization Conference. Three additional
papers describing the other systems are currently in draft mode. See attached documents.

What do you plan to do during the next reporting period to accomplish the goals?

Continue to enhance the different system, continue the validation process, and prepare the
applications for distribution.

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

The four prototype systems that have been designed have generated great interest among
multiple providers, researchers, and administrators. Two senior individuals at the Defense
Health Agency (DHA) have seen the systems and are interested in looking into how we can
integrate some of those tools within the DHA enterprise enclave. In addition, widely recognized
researchers from Johns Hopkins University (JHU) are interested in how to used our visualization
techniques for population health.
What was the impact on other disciplines?

The impact of our work is touching multiple disciplines and research domains including clinical informatics, health IT, computer science, medicine, and population health.

What was the impact on technology transfer?

“Nothing to Report.”

What was the impact on society beyond science and technology?

“Nothing to Report.”

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

“Nothing to Report”

Actual or anticipated problems or delays and actions or plans to resolve them

The project and actions are a little bit behind schedule due to the challenges of finding qualified candidates that can obtain the credentials needed to work within a DoD facility.

Changes that had a significant impact on expenditures

No changes on expenditure.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

“Nothing to Report”

6. PRODUCTS:

- Publications, conference papers, and presentations
  
  Journal publications. Nothing to Report
**Books or other non-periodical, one-time publications.** Nothing to Report

**Other publications, conference papers, and presentations**

- TrajectoryFlow: Visual Summarization of Temporal Sequences [Draft 2019]
- A Reinforcement Learning Approach to Automatically Model and Learn from User Interactions [Draft 2019]

**Website(s) or other Internet site(s)**

“Nothing to Report”

**Technologies or techniques**

The design and development of our different visualization tools have produced novel techniques including:

- Novel graph-based approach to visualize clinical trajectories
- New pixel-based visualization method that works as a look-ahead tool for patients
- Novel sequence modeling algorithm to summarize longitudinal trajectories

1. **Inventions, patent applications, and/or licenses**

“Nothing to Report”

2. **Other Products**

- The four different software tools

7. **PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS**

What individuals have worked on the project?
Name: Jesus Caban, PhD  
Project Role: PI  
Contribution to Project: Dr. Caban has organized meetings, tracked progress of the project, and evaluated various visualization techniques for exploring large clinical data.

Name: Elizabeth Jimenez  
Project Role: Developer  
Contribution to Project: Ms. Jimenez has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Name: Filip Dabek  
Project Role: N/A (Data Scientist for the National Intrepid Center of Excellence)  
Contribution to Project: Mr. Dabek has begun implementing an interface for the visual analytics framework, in addition to developing and evaluating a visualization technique.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?  
“Nothing to Report.”

What other organizations were involved as partners?

8. SPECIAL REPORTING REQUIREMENTS:  
See attachments.  
- W81XWH-15-2-0016 Year 4 Quarter Reports.pdf: copy of all the quarterly reports for year #1.  
- W81XWH-15-2-0016 Year 4 Supplements.pdf: copy of all the papers and draft papers.

8. APPENDICES:  
See attachment.  
- W81XWH-15-2-0016 Year 4 QuadChart.ppt